

NASA TECHNICAL NOTE



NASA TN D-3464

NASA TN D-3464

LOAN COPY: RE
AFWL (WLI)
KIRTLAND AFB,

0130362



TECH LIBRARY KAFB, NM

AN ITERATIVE METHOD FOR COMPUTING THE GENERALIZED INVERSE OF A MATRIX

by H. P. Decell, Jr., and S. W. Kahng

*Manned Spacecraft Center
Houston, Texas*



0130362

NASA TN D-3464

AN ITERATIVE METHOD FOR COMPUTING THE
GENERALIZED INVERSE OF A MATRIX

By H. P. Decell, Jr., and S. W. Kahng

Manned Spacecraft Center
Houston, Texas

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 - Price \$1.00

ABSTRACT

The purpose of this paper is to present extensions of results announced by A. Ben-Israel concerning an iterative method for computing the generalized inverse of an arbitrary complex matrix. Ben-Israel announced his results without proof; at about the same time, the authors of this paper independently derived very similar results with more relaxed hypotheses. These similar results are presented with proof, together with comments pertaining to the Ben-Israel theorem.

AN ITERATIVE METHOD FOR COMPUTING THE GENERALIZED INVERSE OF A MATRIX

By H. P. Decell, Jr., and S. W. Kahng*
Manned Spacecraft Center

SUMMARY

This paper presents extensions of results announced by A. Ben-Israel (ref. 1) concerning an iterative method for computing the generalized inverse of an arbitrary complex matrix. Ben-Israel announced his results without proof; at about the same time, the authors of this paper independently derived very similar results with more relaxed hypotheses. These similar results are presented with proof, together with comments pertaining to the Ben-Israel theorem.

INTRODUCTION

A. Bjerhammar (ref. 2), E. H. Moore (ref. 3), and R. Penrose (ref. 4) independently generalized the concept of matrix inversion to include arbitrary complex matrices. The generalized inverse of a singular or non-square matrix possesses properties that make it a central concept in matrix theory as well as a very useful applied tool in statistical estimation, curve fitting, controllability of linear dynamical systems, stability theory, and so forth (refs. 5 through 14 and 16, 17, and 19).

One of the equivalent definitions of the generalized inverse of an arbitrary complex matrix is an immediate consequence of theorem I due to R. Penrose (ref. 4) stated here without proof.

*

Lockheed Electronics Company at NASA-MSC.

Theorem I (Penrose)

The four matrix equations

$$AXA = A \quad (1)$$

$$XAX = X \quad (2)$$

$$(XA)^* = XA \quad (3)$$

$$(AX)^* = AX \quad (4)$$

have a unique solution X , for each complex matrix A .

The unique solution X in theorem I is denoted $X = A^+$ and is called the generalized inverse of A . In addition, it follows immediately from this definition that if A is square and nonsingular, then A^+ is the usual inverse of A (that is, in classical notation, $A^+ = A^{-1}$).

SYMBOLS

A	a complex matrix
A^*	matrix conjugate transpose of A
A^{-1}	matrix inverse of nonsingular A
A^+	generalized inverse of A
B	a complex matrix
b_{ij}	elements of the matrix B
k, n	positive integers

$P_{R(B)}$	orthogonal projection on range space of B
$R(B)$	range space of B
$R(B)^\perp$	orthogonal complement of $R(B)$
$\text{tr } A$	trace of the matrix A
X	unknown matrix
X_n	matrix iterate
θ	zero matrix or vector as indicated by context
λ_i	eigenvalues
μ, ϵ	vectors
ξ	eigenvector
$\ \cdot \ $	matrix norm

SOME BASIC CONSIDERATIONS

In an effort to make this paper self-contained it is necessary to state some basic lemmas concerning generalized matrix inversion and fundamental matrix theory. Some lemmas will be stated without proof, but with ample reference.

Lemma I

The generalized inverse A^+ of A (as defined by theorem I) is the unique solution of the two matrix equations

$$AX = P_{R(A)} \tag{5}$$

$$XA = P_{R(X)} = P_{R(A^*)} \quad (6)$$

where $R(A)$, $R(X)$, and $R(A^*)$, respectively, denote the range space of A , X , and A^* , and where $P_{R(A)}$, $P_{R(X)}$, and $P_{R(A^*)}$, respectively, denote the orthogonal projection operators on $R(A)$, $R(X)$, and $R(A^*)$.

Proof. - From theorem I it follows that $AX = AA^+$ and $XA = A^+A$ are hermitian idempotent matrices and hence are projection operators on the desired range spaces.

Lemma II

For the generalized inverse A^+ of A (ref. 4)

$$A^+AA^* = A^* = A^*AA^+ \quad (7)$$

$$(A^+)^+ = A \quad (8)$$

$$(A^*)^+ = (A^+)^* \quad (9)$$

Definition I

The norm of a square matrix B is a non-negative number (denoted $||B||$) which satisfies (ref. 15).

$$||B|| > 0, \text{ if } B \neq \theta, \quad ||\theta|| = 0 \quad (10)$$

$$||cB|| = |c| \cdot ||B||, \text{ for any complex number } c \quad (11)$$

$$\|B + D\| \leq \|B\| + \|D\| \quad (12)$$

$$\|BD\| \leq \|B\| \cdot \|D\| \quad (13)$$

Definition II

A sequence of square matrices B_n ($n = 0, 1, 2, \dots$) is said to converge in norm $\|\cdot\|$ to the matrix B (written $B_n \xrightarrow{\|\cdot\|} B$) provided that the sequence of real numbers $\|B_n - B\|$ converges to zero.

Following (ref. 15), a number of basic norms are defined, and some comparison inequalities are stated.

Lemma III

If B is a k by k square matrix then the following equalities define norms satisfying definition I (ref. 15)

$$\|B\| = \sqrt{\max_i \lambda_i} \quad (14)$$

where $\lambda_1, \dots, \lambda_k$ are the eigenvalues of B^*B (see lemma V).

$$\|B\|_1 = \max_i \sum_{j=1}^k |b_{ij}| \quad (15)$$

$$\|B\|_2 = \max_j \sum_{i=1}^k |b_{ij}| \quad (16)$$

$$\|B\|_3 = k \max_{i,j} |b_{ij}| \quad (17)$$

$$\|B\|_4 = \sqrt{\sum_i \sum_j |b_{ij}|^2} = \sqrt{\text{tr } B^* B} \quad (18)$$

Lemma IV

The norms in lemma III satisfy the following inequality (ref. 15)

$$\frac{1}{k} \|B\|_1 \leq \|B\|_2 \leq \frac{1}{k} \|B\|_3 \leq \|B\|_4 \leq \sqrt{k} \|B\| \quad (19)$$

Lemma V

For any matrix A , the eigenvalues of AA^* and A^*A are real, identical, and non-negative.

ITERATIVE COMPUTATION OF A^+

Ben-Israel initially published the following theorem (ref. 18) in an attempt to give a useful iterative scheme for computing A^+ . This theorem used the equivalent definition of A^+ given in lemma I.

Theorem II (Ben-Israel)

The sequence of matrices defined by

$$X_{n+1} = X_n (2P_{R(A)} - AX_n)$$

converges in any $\|\cdot\|$ -norm defined by lemma III to the generalized inverse of A , provided

$$X_0 = A^* B_0, \text{ for some nonsingular } B_0 \quad (20)$$

$$X_0 = C_0 A^*, \text{ for some nonsingular } C_0 \quad (21)$$

$$\| \| P_{R(A)} - AX_0 \| \| < 1 \quad (22)$$

$$\| \| P_{R(A^*)} - X_0 A \| \| < 1 \quad (23)$$

Note here that the term "the sequence X_n converges in $\| \| \|$ -norm to the generalized inverse of A " means that (since X_n may be a rectangular matrix for which a norm, in our sense, is not defined)

$$(P_{R(A)} - AX_n) \xrightarrow{\| \| \|} \theta \quad (24)$$

$$(P_{R(A^*)} - X_n A) \xrightarrow{\| \| \|} \theta \quad (25)$$

Theorem II requires a good deal of hypotheses. In particular, it requires a priori knowledge of the projections $P_{R(A)}$ and $P_{R(A^*)}$. However, as Ben-Israel notes, a few more direct computations will produce the generalized inverse without iteration in this particular case.

During preparation for the publication of these results, Ben-Israel announced without proof very similar results (ref. 1). In fact, the statement of the main result in theorem III will closely parallel the statement of the Ben-Israel results announced in reference 1. A corollary will follow theorem III in order to point out results apparently unnoticed by Ben-Israel.

Theorem III

Let A be a q by p matrix (nonzero), λ_m be the largest eigenvalue of AA^* , and $X_0 = \alpha A^*$, where $0 < \alpha < 1/\lambda_m$. The sequence defined by

$$X_{n+1} = X_n (2I - AX_n) \quad (n = 0, 1, 2, \dots)$$

converges in any $|| ||$ -norm defined in lemma III to the generalized inverse of A .

Again, note that this convergence is that convergence defined by equations (24) and (25).

Proof.- The following facts will be established in order to prove the desired results for $X_0 = \alpha A^*$ satisfying the hypothesis

$$P_{R(A)} - AX_{n+1} = (P_{R(A)} - AX_n)^2 \quad (26)$$

$$P_{R(A^*)} - X_{n+1}A = (P_{R(A^*)} - X_nA)^2 \quad (27)$$

$$|| P_{R(A)} - AX_0 || < 1 \quad (28)$$

$$|| P_{R(A^*)} - X_0A || < 1 \quad (29)$$

The indicated norm in equations (28) and (29) will be the square root of the largest eigenvalue of B^*B defined in lemma III. Once these facts are established the proof will be complete since equations (26) through (29), together

with equation (13) of definition I and lemma IV, imply the convergence (in any norm of lemma III) of $P_{R(A)} - AX_n$ and $P_{R(A^*)} - X_n A$ to θ .

In order to establish equation (26) note, from the recursive computation of X_n , that there exist matrices F_n , ($n = 0, 1, 2, \dots$) such that

$$X_n = F_n A^* \quad (30)$$

so that, using equations (1), (5), (7), and (30),

$$AX_n P_{R(A)} = AF_n A^* AA^+ = AF_n A^* = AX_n \quad (31)$$

and

$$P_{R(A)} AX_n = AA^+ AX_n = AX_n \quad (32)$$

From equations (31) and (32), observe that

$$\begin{aligned} \left(P_{R(A)} - AX_n \right)^2 &= P_{R(A)} - 2AX_n + AX_n AX_n \\ &= P_{R(A)} - AX_n (2I - AX_n) \\ &= P_{R(A)} - AX_{n+1} \end{aligned}$$

A dual argument will establish equation (27).

In order to establish equation (28), first note that if λ_i is an eigenvalue of AA^* then $1 - \alpha\lambda_i$ is an eigenvalue of $I - \alpha AA^*$. This fact will be of importance in the examination of the eigenvalues of $AA^+ - \alpha AA^*$ leading to the proof of equation (28). However, it will first be necessary to prove that the nonzero eigenvalues of $AA^+ - \alpha AA^*$ are of the form $1 - \alpha\lambda_i$ where λ_i is a nonzero eigenvalue of AA^* . To this end, let $\xi \neq \theta$ be an eigenvector of $AA^+ - \alpha AA^*$ with the associated eigenvalue λ_ξ . The vector $\xi \neq \theta$ can be written as the sum $\xi = \mu + \eta$, where $\mu \in R(A)$ and $\eta \in R(A)^\perp$ (ref. 5) and hence it follows that

$$(AA^+ - \alpha AA^*)(\mu + \eta) = \lambda_\xi(\mu + \eta) \quad (33)$$

$$(AA^+ - \alpha AA^* AA^+)(\mu + \eta) = (AA^+ - \alpha AA^*)(\mu + \eta) \quad (34)$$

so that by lemma I

$$\mu - \alpha AA^* \mu = \lambda_\xi \mu + \lambda_\xi \eta \quad (35)$$

Multiplying both sides of equation (35) by AA^+ (and using eqs. (1) and (5)) it follows that

$$\mu - \alpha AA^* \mu = \lambda_\xi \mu \quad (36)$$

$$(I - \alpha AA^*)\mu = \lambda_\xi \mu \quad (37)$$

Now if $\mu \neq \theta$, then equation (37) implies that λ_ξ is also an eigenvalue of $(I - \alpha AA^*)$. In this case, as mentioned previously, λ_ξ must be of the form

$$\lambda_\xi = 1 - \alpha\lambda_i \quad (38)$$

for some eigenvalue λ_i of AA^* . Moreover, for the case $\mu \neq \theta$, it will be shown that the λ_i in equation (38) are different from zero. To this end, note that if $\lambda_i = 0$ and $\mu \neq \theta$ then $\lambda_\xi = 1$ so that equation (36) implies

$$AA^* \mu = \theta \quad (39)$$

Multiplying both sides of equation (39) by $A^+ A^+$ it follows that

$$A^+ A^+ AA^* \mu = A^+ A^* \mu = (AA^+)^* \mu = AA^+ \mu = \theta \quad (40)$$

This is impossible since $\mu \neq \theta$ and $\mu \in R(A)$. Indeed, AA^+ is the orthogonal projection on the range of A so that

$$AA^+ \mu = \mu \neq \theta \quad (41)$$

contrary to equation (40).

Considering the case $\mu = \theta$, it follows from equation (35) that $\lambda_\xi = 0$. Hence, the nonzero eigenvalues of $AA^+ - \alpha AA^*$ are of the form $1 - \alpha\lambda_i$ where λ_i is a nonzero eigenvalue of AA^* . Moreover, all of the eigenvalues

AA^* are non-negative so that for $0 < \alpha < 1/\lambda_m$ where $(\lambda_m \neq 0$ is the largest eigenvalue of AA^*)

$$1 - \alpha\lambda_i < 1 \quad (42)$$

Since, as mentioned at the outset of the proof,
 $||P_{R(A)} - \alpha AA^*|| = ||AA^+ - \alpha AA^*||$ is the square root of the largest eigenvalue of

$$(AA^+ - \alpha AA^*)(AA^+ - \alpha AA^*)^* = (AA^+ - \alpha AA^*)^2 \quad (43)$$

it follows that the nonzero eigenvalues of $(AA^+ - \alpha AA^*)^2$ are of the form

$$(1 - \alpha\lambda_i)^2 < 1 \quad (44)$$

so that

$$||AA^+ - \alpha AA^*|| = \sqrt{\max_i \left\{ (1 - \alpha\lambda_i)^2, 0 \right\}}$$

where the λ_i are nonzero eigenvalues of AA^* . Hence

$$||AA^+ - \alpha AA^*|| < 1$$

which is in fact equation (28).

A dual argument will establish equation (29) (using the fact that AA^* and A^*A have the same eigenvalues) and, thus, complete the proof of the theorem.

The following corollary will eliminate the need for computing the eigenvalues of AA^* in theorem III.

Corollary I

In theorem III, the choice of α may be limited to $0 < \alpha < 1/\beta$ where β is any norm defined in lemma III of AA^* .

Proof - The eigenvalues of any square matrix B cannot exceed any of the norm of B defined in lemma III (ref. 15). Hence, if β is any norm of AA^* defined by lemma III then

$$0 < \lambda_m \leq \beta$$

so that

$$0 < \frac{1}{\beta} \leq \frac{1}{\lambda_m}$$

It follows that theorem III is valid for any choice of α such that

$$0 < \alpha < \frac{1}{\beta}$$

CONCLUDING REMARKS

The algorithm described in theorem III always guarantees the knowledge of the proper initial guess to force convergence. In addition, corollary I rules out the necessity for calculating eigenvalues of the matrix AA^* in order to find a suitable constant α . This item did not appear in reference 1. The algorithm in theorem III was successfully tested on Hilbert segments through order seven.

Manned Spacecraft Center
National Aeronautics and Space Administration
Houston, Texas, May 2, 1966

REFERENCES

1. Ben-Israel, A. : An Iterative Method for Computing the Generalized Inverse of an Arbitrary Matrix. Notices of the Amer. Math. Soc., vol. 12, no. 6, Oct. 1965, p. 708.
2. Bjerhammar, A. : Application of Calculus of Matrices to the Method of Least Squares With Special Reference to Geodetic Calculations. Trans. Roy. Inst. Tech. Stockholm, vol. 49, 1951, pp. 1-86.
3. Moore, E. H. : Bull. Amer. Math. Soc., vol. 26, 1920, pp. 394-395.
4. Penrose, R. : A Generalized Inverse for Matrices. Proc. Camb. Philos. Soc., vol. 51, 1955, pp. 406-413.
5. Charnes, A. ; and Ben-Israel, A. : Contributions to the Theory of Generalized Inverses. J. Soc. Indust. Appl. Math., 11(1963), pp. 667-699.
6. Cline, Randal E. : Note on the Generalized Inverse of the Product of Matrices. SIAM Rev., vol. 6, 1964, pp. 57-58.
7. Cline, R. E. : Representations for the Generalized Inverse of Matrices With Applications in Linear Programming. Doctoral Thesis, Purdue Univ., 1963.

8. Cline, R. : On the Computation of the Generalized Inverse A , of an Arbitrary Matrix A , and the Use of Certain Associated Eigenvectors in Solving the Allocation Problem. Preliminary Report, Statistical and Computing Laboratory, 7, Purdue Univ., 1958.
9. Decell, H. P. : A Special Form of the Generalized Inverse of an Arbitrary Complex Matrix. NASA TN D-2748, 1965.
10. Decell, H. P. : An Application of the Cayley-Hamilton Theorem to Generalized Matrix Inversion. SIAM Rev., vol. 7, no. 4, Oct. 1965, pp. 526-528.
11. Decell, H. P. : An Alternate Form of the Generalized Inverse of an Arbitrary Complex Matrix. SIAM Rev., vol. 7, no. 3, July 1965, pp. 356-358.
12. Decell, H. P. ; and Odell, P. : A Note Concerning a Generalization of the Gauss-Markov Theorem. Journal of the Texas Academy of Science, March 1965.
13. Decell, H. P. : An Application of Generalized Matrix Inversion to Sequential Least Squares Parameter Estimation. NASA TN D-2830, May 1965.
14. Desoer, C. A. ; and Whalen, B. H. : A Note on Pseudoinverses. J. Soc. Indust. Appl. Math., vol. 11, no. 2, June 1963, pp. 442-447.
15. Faddeev, D. K. ; and Faddeeva, V. N. : Computational Methods of Linear Algebra. W. H. Freeman Co., San Francisco, 1963, pp. 260-265.
16. Greville, T. N. E. : The Pseudoinverse of a Rectangular or Singular Matrix and the Application to the Solution of Systems of Linear Equations. SIAM Rev., vol. 1, 1959, pp. 38-43.
17. Kalman, R. E. ; Ho, Y. C. ; and Narendra, L. S. : Controllability of Linear Dynamical Systems. Contributions to Differential Equations, vol. I, Interscience Publishers, Inc., New York, 1962.
18. Ben-Israel, A. : An Iterative Method for Computing the Generalized Inverse of an Arbitrary Matrix. Mathematics of Computation, vol. 19, no. 91, July 1965, pp. 452-455.
19. Price, C. : The Matrix Pseudoinverse and Minimal Variance Estimates. SIAM Rev., 6(1964), pp. 115-120.

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Technical information generated in connection with a NASA contract or grant and released under NASA auspices.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

TECHNICAL REPRINTS: Information derived from NASA activities and initially published in the form of journal articles.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities but not necessarily reporting the results of individual NASA-programmed scientific efforts. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C. 20546